图书情报工作
第62卷 第3期 2018年2月

基于自引网络和主路径分析的论文主题创新实证研究*

■ 魏瑞斌
安徽财经大学管理科学与工程学院 蚌埠 233030

摘要：[**目的/意义**]探讨如何从大量论文中快速发现相关论文，研究论文主题创新的方法，为研究者阅读和利用期刊论文提供参考。[**方法/过程**]在对论文主题创新和自引网络概念界定的基础上，提出研究论文主题创新的思路、方法和工具。以美国印第安纳大学在图书情报学领域的论文为例进行实证研究。[**结果/结论**]利用主路径分析可以快速从自引网络中发现由于引用而产生联系的论文，在此基础上可以进行主题创新分析。

关键词：主题创新　主路径分析　自引网络　研究主题
分类号：G250
DOI：10.13266/j.issn.0252-3116.2018.03.008

1 引言

创新是科学共同体获得社会承认的根本依据。把创新作为科学共同体的行为规范，是要求“科学研究成果总应该是新颖的。一项研究没有给充分了解和理解的东西增添新内容，则无所贡献于科学”。科研人员通过生产知识可以获得科学共同体和社会的承认与奖赏，而创新成为科学家获得社会承认的根本依据[1]。期刊论文是科研人员创新成果的一种重要表现形式。通过期刊论文的内容可以了解其创新之处。胡英奎等[2]、盛杰[3]、朱大明[4]从学术期刊编辑的角度，探讨了评价论文创新性的一些具体做法和科技期刊论文创新性鉴审的4个基本要素。曹妍等[5]利用德尔菲法构建了护理论文创新性评价指标体系。这些研究对于评价单篇论文的创新性提供了较为具体和科学的参考。但是，对于已经发表过的论文，后续研究者如何快速发现其研究内容的创新，这些方法就有较大的局限性。本研究尝试利用主路径分析方法和内容分析方法相结合的方式来对论文研究主题的创新进行探索性研究。由于创新的相对性特点，本研究将相关研究界定在一个机构的研究成果范围之内进行探讨。

2 论文主题创新与自引网络和主路径分析

2.1 论文主题创新

论文创新可以有不同角度，如研究主题、研究方法或研究结论等。本研究探讨的是期刊论文的主题创新。在判断期刊论文主题是否有创新时，通常是将其与已有的研究成果比较之后得出的结果。判断论文主题创新可以分为3个层面的内容，一是有没有创新；第二是在哪些方面有创新；第三是创新的程度。本研究针对第一个和第二个层面进行研究，即论文的主题是否具有创新性。

论文主题创新是对一个特定数据集中的论文主题进行比较、分析的结果。在判断论文主题创新的时候需要考虑很多因素：①是时间因素，在判断论文主题创新时通常是把某篇论文与其之前的论文进行比较，分析其主题是否具有创新，分析的依据通常是反映其主题的专业术语。②范围因素，如以一个学者的期刊论文为数据集，可以探讨其个人的科研创新；以一个机构在特定领域的研究成果为数据集，可以研究该数据集范围内该机构研究人员的创新。本研究的研究对象主要局限在一个机构范围之内。机构论文主题创新的判断有两个角度，一个是对机构内研究者的研究成果进

* 本文系国家社会科学基金项目“期刊论文创新性评价标准及其计量模型研究”（项目编号：12BTQ034）研究成果之一。
作者简介：魏瑞斌（ORCID：0000-0001-6271-7881），教授，博士，E-mail：rbwxy@126.com。
收稿日期：2017-07-24　修回日期：2017-11-20　本文起止页码：64-70　本文责任编辑：王传清

行比较,一个是放在一个更大的范围内与机构外的研究成果进行比较。本研究侧重于第一种情况的比较。③判断论文主题创新通常是根据反映其研究主题的专业术语,如A论文的研究主题是共词分析,B论文的研究主题是信息检索。这些专业术语可以来自论文的题名、摘要、关键词或全文等途径,如T. Amajd等[6]在对作者、期刊等进行排序时,将其分为多媒体检索、医学信息检索和数据库和查询处理3个研究主题。

2.2 自引网络

通过论文引用的关系,可以发现不同研究主题之间的关系,也可以作为判断论文研究主题创新的一个角度。I. Hellsten等[7]提出了作者自引网络,并指出作者自引网络能够更好地发现作者新的研究主题。J. Y. Lee[8]认为,作者自引网络能够成功地识别作者的核心论文和领先成果。笔者认为,自引网络可以分为不同层次。作者自引网络[9]是指某作者及其合作者撰写的论文集合中,论文之间引用与被引所形成的引文网络。机构自引网络则是同一机构的研究者,引用自己的文献或者是引用同一机构内其他研究者的成果,而形成的引文网络。在本研究的机构自引网络中,节点代表机构研究者发表的论文,节点之间的连线表示论文之间的引用关系。自引网络还可以包括更加宏观的地区自引网络和国家自引网络等。

通常情况下,一个研究者与同一机构同一学科的其他研究者之间由于空间等因素,相互之间的学术交流比较频繁,对彼此的研究主题较为了解,合作的可能性比较大。机构自引网络一方面反映出同一机构内的研究者,在从事科学研究过程中,其关注的知识在自己和同一研究机构研究者之间的传播过程,同时相对客观地反映了同一机构研究者的研究主题的变化。

2.3 主路径分析

主路径分析最早是由N. Hummon和P. Dereian[10]从网络连通性出发提出的,其主要目标是通过识别出引文网络中具有最大连通度的系列文献来概述研究领域的发展态势以及领域演化过程中的主要文献、主要人物与主要事件。W. Goffman[11]、M. Jahn等[12]基于引文网络方法的研究结果表明:一个专业是由其历史发展中出现的为数不多且极其重要的事或人所定义的。这一结论为引文网络主路径方法的产生提供了理论支持。

主路径分析的理论前提是将引文网络看作一个输送知识信息的渠道系统。如果一篇论文能够把之前一些论文的知识整合到一起,并且为新知识的增长做出实质性贡献,那么这篇论文就有可能被大量引用,而且有可能使此后再引用此前的论文变得有点多余。因此,这种论文就成了渠道系统中的重要枢纽,大量知识信息从此处流过[13]。韩毅和金碧辉[14]深入分析了引文网络主路径分析方法的产生背景、基本内涵与算法实现,并总结主路径分析在理论及应用研究中存在的主要问题。隗玲和方曙[15]研究发现,学者们对主路径方法已有的修正和拓展工作主要聚焦于主路径方法的选择原则、搜寻起点的确定和弧的权重设置3个部分。

从宋歌[16]、韩毅[17]和祝清松[18]等研究看,主路径分析方法可以帮助研究者快速从引文网络中发现重要的研究成果,并且能直观反映出研究成果之间的引用和被引关系。本研究利用的主路径分析方法是将论文通过引用关系连接在一起,同时结合时间因素来分析这些论文在研究主题方面的创新。

3 研究思路、研究工具与方法

3.1 研究思路

3.1.1 收集数据 以某个机构名称为检索对象,从特定引文数据库中获取其论文数据。由于数据格式和处理工具的局限,本研究使用采集自Web of Knowledge平台的期刊论文数据。

3.1.2 数据处理 从Web of Knowledge收集的数据,利用HistCite软件的Graph Maker功能可以快速生成数据对应的自引网络。自引网络的数据可以直接保存成.net文档。同时,利用HistCite软件还可以自动统计出每篇论文的局部被引、全局被引等相关数据,为后续的数据分析提供数据支持。.net文档可以直接导入Pajek软件中做进一步的处理。在Pajek中可以直接获取每篇论文的点度中心度等网络计量指标,并可以生成可视化的引文网络图。

3.1.3 数据分析 结合利用Pajek生成的引文网络图和每篇论文的相关数据,可以进一步分析论文在整个引文网络中所处的位置;通过对论文原始信息分析可以确定其研究主题。结合论文的时间属性来分析机构研究主题创新。

3.2 研究方法与研究工具

本研究应用主路径分析方法(main path analysis,MPA)对自引网络进行研究。主路径分析方法是一种用于分析时间流的特殊技术。引文网络可以看作是一个输送科学知识或信息的渠道系统。如何知识通过引文关系而流通,那么参与许多论文之间路径的某个引文关系,就要比很少参与论文之间路径的另一个引文关系重要。那些最重要的引文关系就形成了一条或多

条主路径，这可能是一项研究传统的骨架[13]。

Pajek 4.06 版本与以前版本的功能有一定差异，新版本中提供了主路径分析的不同方法供用户选择。如：①Network -Acylic Network -Create Weighted Network + Vector -Traversal Weights -Search Path Count (SPC)/ Search Path Link Count(SPLC)/ Search Path Node Count (SPNP)；②*Network-Acylic Network -Creat* (*Sub*) *Network-Critical Path Method-CPM*。隗玲和方曙[11]的研究中对这些不同的方法进行了具体的解释。本研究利用第 2 种方法来寻找引文网络中的主路径。

4 实证研究

4.1 数据获取

本研究以 Web of Science 核心合集为数据源，选取印第安纳大学学者在信息科学和图书馆学领域 1986 年 –2017 年发表的论文。最终检索结果为 731 条记录。具体检索式如下：

地址：(Indiana University)

精炼依据：Web of Science 类别：(INFORMATION SCIENCE LIBRARY SCIENCE)

时间跨度：1986 – 2017。索引：SCI-EXPANDED，CPCI-S，CPCI-SSH，CCR-EXPANDED，IC。

从图 1 中可以看出，1990 年 – 1996 年，机构的发文数量呈现为一个快速增长的趋势；1997 年开始有一个下降，但其整体保持在 20 篇 – 40 篇之间，少数年份达到了 50 篇。这从一定程度上反映出印第安纳大学在信息科学和图书馆学领域的论文数量较为稳定。

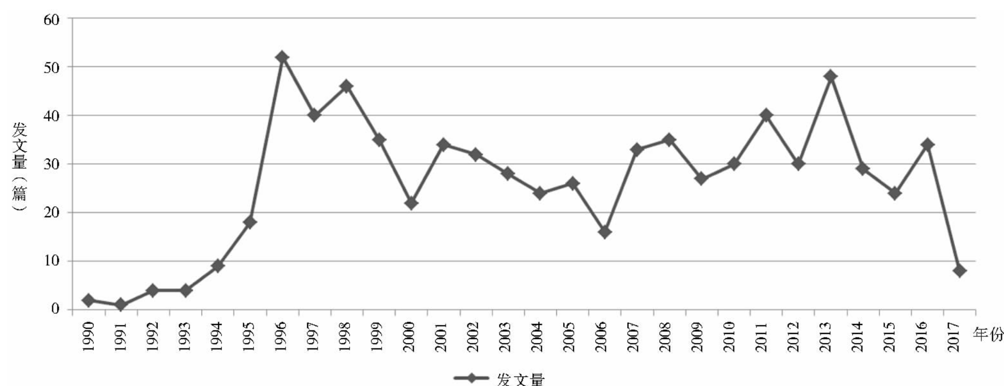


图 1 印第安纳大学在信息科学和图书馆学领域发文数量分布

本研究首先使用 Histcite 作为研究工具，自动生成了 731 篇论文的引文网络，并将引证网络原始数据保存为 .net 文档，再导入 Pajek 进行后期处理。

4.2 数据分析

4.2.1 引证网络整体分析 点度是指一个节点所拥有的连线数量，它是一种离散属性。在引证网络中，它表征一篇论文引用其它论文或被其它论文引用的数量。Pajek 的分区功能(partition) ，可以计算出每个节点的度数(包括引用网络中其它论文的数量和被其它论文引用的数量) ，也可以单独计算。在整个引证网络中，有 731 个节点，497 条连线，其密度为 0.000 9，节点平均度为 1.359 8。这些数据反映出此引证网络中论文之间的引用关系并不是特别密切。从表 1 可以看出，点度为 0 的节点有 382 个，占总体数量的 52.26%，即有一半以上的论文与其它论文之间没有引用和被引用关系；有 133 篇论文与其它论文之间的引用和被引只有 1 次。这反映出印第安纳大学的学者在信息科学和图书馆学领域的引用行为存在着集中与分散的现象。

从统计结果看，剔除孤立节点之后的部分规模较小的连通子网络中，节点数量大部分在 2 个到 6 个之间，但有一个包含 238 个节点的最大连通子网络。本研究将最大连通子网络中的 238 篇论文按其发表时间分为 5 个时间窗口：1992 年 – 1997 年、1998 年 – 2002 年、2003 年 – 2007 年、2008 年 – 2012 年和 2013 年 – 2017 年。然后利用 Pajek 绘制 5 个时间窗口内论文间引证网络的变化情况。从图 2 可以看出，第 1 个时间窗口只有 10 篇论文，其中有 2 篇有引用关系，另外 8 篇都是孤立的节点；第 2 个时间窗口有 56 个节点，这 56 个节点已经形成了一个较大的连通网络；第 3 个到第 5 个时间窗口的节点数分别为 111 个、181 个和 238 个。随着时间的推移，论文间由于引用关系，越来越多的论文节点被连接起来，最终形成一个较大的连通网络。从论文引用的角度看，连通子网络反映了其研究主题之间的延续性，而孤立节点则反映其研究主题之间的分散性。

表 1　网络中各个节点的点度分布情况

Cluster	Freq	Freq(%)	CumFreq	CumFreq(%)	Representative
0	382	52.257 2	382	52.257 2	1 PUTTAPITHAKPORN S, 1990
1	133	18.194 3	515	70.451 4	4 NISONGER TE, 1992
2	93	12.722 3	608	83.173 7	43 Rosenbaum H, 1996
3	36	4.924 8	644	88.098 5	37 OVERHAGE JM, 1995
4	29	3.967 2	673	92.065 7	159 Harter SP, 1998
5	14	1.915 2	687	93.980 8	226 Mostafa J, 2000
6	9	1.231 2	696	95.212	198 Cronin B, 1999
7	7	0.957 6	703	96.169 6	175 Mostafa J, 1998
8	8	1.094 4	711	97.264	119 Cronin B, 1997
9	9	1.231 2	720	98.495 2	176 Cronin B, 1998
10	1	0.136 8	721	98.632	544 Yan EJ, 2011
11	4	0.547 2	725	99.179 2	259 Cronin B, 2001
12	3	0.410 4	728	99.589 6	233 Kling R, 2000
13	1	0.136 8	729	99.726 4	306 Borner K, 2003
14	1	0.136 8	730	99.863 2	526 Ding Y, 2011
16	1	0.136 8	731	100	578 Lariviere V, 2012

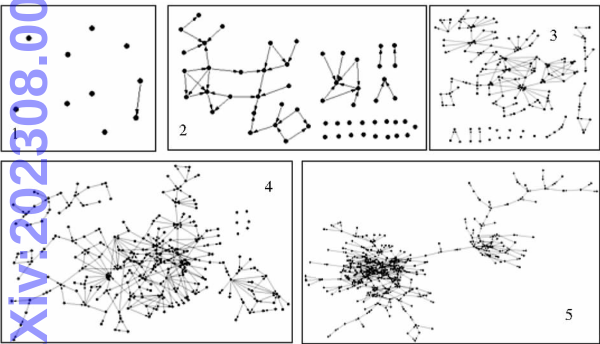


图 2　机构自引网络最大子网络的变化

本研究根据节点的点入中心度和点出中心度,将 238 个节点分为 3 种类型。

(1)知识输出型论文。此类型节点的点出中心度大于点入中心度(两者差值大于等于 3),约占总体数量的 14%。这些论文是该机构某个研究主题的早期研究成果,对后续相关研究有引领作用。如 41 号论文[19],被网络中其它 9 篇论文引用,但是没有引用其它论文。该文提出了一个解释决策支持系统功能绩效的新模型。

(2)知识吸收型论文。此类型节点的点入中心度大于点出中心度(两者差值大于等于 3),约占总体数量的 13%。此类论文通常是在对前期比较多的研究成果梳理的基础上,形成的较新的阶段性研究成果。如第 148 号论文[20],引用了网络中的 10 篇论文,但只被其它论文引用 1 次。该文在社会临场感理论等基础上提出了一个合作研究的整合理论模型。

(3)知识吸收与输出均衡论文。此类型节点的点入中心度与点出中心度相差不大(两者差值小于等于 2),约占总体数量的 73%。这类论文是机构中间阶段的研究成果。如 126 号论文[21],引用网络中的 3 篇论文,被网络中其它 5 篇论文引用。该文是作者对他在 1999 年提出的媒介同步性理论(media synchronicity theory (MST))的进一步完善和拓展。

从整个引证网络的视角看,在被引频次接近的情况下,知识输出型论文的创新性较高,其次是知识吸收型论文;第三是知识输入与输出均衡型论文。依据这个标准对论文创新程度进行比较时,需要选择同一个时间窗口的同类论文进行比较。如知识输出型论文发表的时间较早,其被引高可能来源于时间的累积优势。单独通过被引频次来比较知识输出型论文和知识吸收型论文的创新是不合理的。

4.2.2　主路径分析　当节点数较少时,可以通过观察网络结构中节点所处的位置来判断一篇论文在引证网络中的地位和作用。但当节点数较多的时候,快速从网络中寻找一些重要的论文难度就会加大。本研究首先利用 Pajek 对 238 篇论文进行处理,结果见图 3(图 3 中中文内容为手工添加)。根据论文内容,本研究将图 3 中的期刊论文分为以下 6 个方面。

(1)电子期刊与学术交流。该方面的 4 篇论文发表于 1996 年-2000 年期间,代表性人物是 S. P. Harter。文献 50 和文献 159 从期刊论文参考文献等角度研究了电子期刊对学术交流的影响。文献 197 和文献

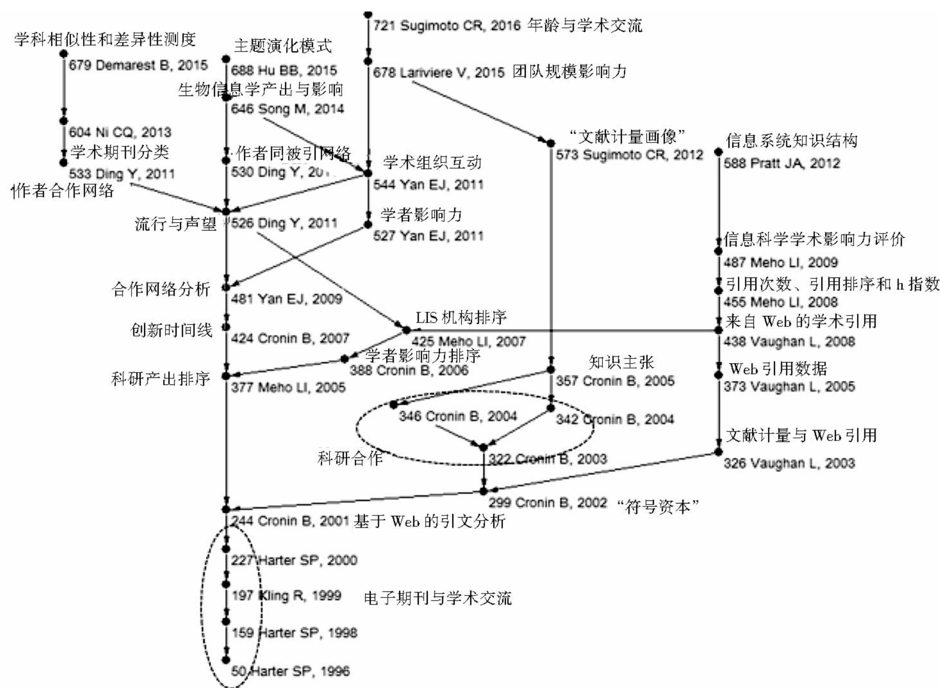


图 3　基于主路径分析的引证网络

227 研究了学术交流中电子期刊出版的政策和实践。结合该数据集整个引文网络看,文献 50、159、197 和 227 被引的次数分别为 2、3、6 和 2 次。这反映出该机构在这个领域还有一系列相关成果,该方法只呈现出了少部分的论文。另外,文献 159 和文献 197 的全局被引分别达到 53 次和 72 次,这反映出这两篇文献同时也得到了机构外学术同行的认可。

(2) Web 引用。在 2001 年 – 2008 年期间,网络引用成为印第安纳部分学者关注的一个研究主题。其代表人物是 B. Cronin 和 L. Vaughan。从图 3 可以看出,文献 244 在网络中占据了一个较为重要的地位,它对网络中其它相关研究起到了比较重要的引领作用。B. Cronin[22] 认为,基于 Web 的引文分析为文献计量学领域带来了新的机遇。该文既是对其机构内其他同行相关研究的一个总结性延续,同时也成为后续相关研究的一个基础。该文在网络结构中起着一个承前启后的重要作用。B. Cronin[23] 的论文标题中使用了符号资本的概念,但其实际上是利用引文次数、web 点击率和媒体提及率 3 个指标对 25 位学者进行相关性分析。文献 326、373 和 438 中,L. Vaughan 和 D. Shaw[24] 对 4 个学科的 web 引用数据和来自 Web of Science (WoS) 等途径的引文数据进行了相关研究。这些研究是目前图书情报领域比较热门的"替代计量学"的研究内容。从余厚强和邱均平[25] 的研究看,替代计量学的概念是 Priem 在 2010 年提出的。B. Cronin 等人自 2001 年起就展开了相关研究,也体现了其研究的创新性和前瞻性。

(3) 科研合作。在 2003 年 – 2005 年期间,利用文献信息研究科研合作成为其研究主题。代表人物是 B. Cronin。如 B. Cronin 和 D. Shaw 等利用传统的文献计量学方法对心理学和哲学领域的科研合作、20 世纪化学领域的合作模式、科研合作对学术写作的影响等进行了相关研究。V. Lariviere 等[26] 通过作者数量、地址数量和国家数量 3 个指标对团队规模对学术影响力的影响进行了研究,发现团队规模越大,作者分布越广泛,其论文获得的被引次数越多。

(4) 学术影响力评价。在 2005 – 2009 年期间,以 L. Meho 为代表的学者等利用传统文献计量学指标和方法对科研人员和机构的学术影响力进行了相关研究。如 L. Meho 对 LIS 的科研人员和机构的科研产出进行研究、利用 h 指数分析信息科学家的学术影响力、基于 Web of Science、Scopus 和 Google scholar 3 个数据源对 25 个 LIS 领域的科研人员的学术影响力进行分析;基于 Scopus 和 Web of Science 数据对随机挑选的 80 位 LIS 研究人员进行了相关研究。从论文引用的角度来看,该文对后续相关研究也产生了较大的影响。

(5) 学术网络分析及应用。自 2009 年开始,各类学术网络的研究成为该机构的一个研究主题。代表人物是丁颖和晏乐伽等。他们从网络视角对各类学术网络进行了相关研究。如利用网络中心性指标对合作网

络进行分析;利用 PageRank 算法对作者影响力进行分析;基于引文网络和合作网络对学术组织互动进行分析;利用有权值的 PageRank 算法对作者引证网络测度学者声誉和影响力进行分析;基于主题的 PageRank 算法对作者引文网络等展开了相关研究。与传统的文献计量学研究相比,这类研究充分利用了社会网络分析、复杂网络和计算机科学等领域的研究成果与文献网络有机结合,是文献计量学领域一个较新的研究方向。

(6)其它方面。在整个网络中,还有一些较为特别的研究主题。如 J. A. Pratt 等[27]针对信息管理系统 25 种学术期刊的数据,利用共引分析、多维尺度分析和主成分分析等方法研究了该领域的知识结构。从论文内容看,J. A. Pratt 等在数据选择时参考了文献 487 的研究内容。两篇研究主题虽然有一定差别,但在数据来源方面有一定的共性。C. R. Sugimoto 和 B. Cronin[28]通过作者风格、学术产出效率与模式、合作模式等信息对 6 位杰出的文献计量学家进行了"文献计量画像"。C. R. Sugimoto 等[29]对社会学、经济学和政治学的 1 000 多位学者的年龄与其科研产出、合作和影响力之间关系进行了定量分析。这些成果都是以论文原始信息和被引信息为基础,针对不同的研究问题进行相关研究。

总体看来,前面 5 个研究主题在一定时间内比较稳定,侧重于对一个研究主题深度的创新。而随着时间变化,研究主题在不断改变,反映其在研究广度上的创新。从图 3 来看,文献 50、159、197 和 227 这 4 篇论文发表于 1996 年 – 2000 年期间,这反映出研究者对该主题研究的持续性。再如文献 388 和 527 研究的都是学者影响力的问题,但是所采用的研究方法发生了变化,这体现出研究主题虽然稳定,但研究内容方面都有创新。文献 244[30]是一篇对 web 引文分析的研究成果,通过文献引用关系,连接了电子期刊与学术交流和 web 引用两个研究主题。这反映了两个研究主题之间有一定的关联性。

5 结论

本研究尝试利用主路径分析方法,快速从一个比较大的网络中提取了一部分引用关系更为密切的文献,通过这些文献的内容分析来对论文主题创新进行比较。这种方法还可以应用在更微观的学者引文网络或更宏观的地区、国家引文网络的研究。这种方法由于只是基于主路径分析方法快速发现引用关系比较紧密的一些文献,因此并不能反映某机构论文主题的全部内容。如图 3 中出现的文献基本都可以归结为科学计量学的范畴,但从原始数据可以发现,印第安纳大学学者还涉及到信息系统、社会信息学等很多方面的研究主题。如果一个机构的论文主题比较分散或研究成果数量较少,或者机构内研究者之间的相互引用或自引较少时,这种方法都有可能不适用。机构自引网络构建时,对于没有引用关系的文献,在网络中都表现为独立的节点,在利用主路径分析时,这些独立的节点都会被排除。但机构自引网络可以把机构研究相关性较强的文献的关系通过引用关系揭示出来,对于研究论文的创新有很好的支撑。

论文主题创新的分析依赖于对论文原始信息的解读,对于少量文献,采取人工方式有一定可行性,但如果是对大量文献的处理,则需要借助于一些更为有效的自动化、智能化处理手段。人工方式有较强的主观性,单纯用一个词语来概括论文的内容不能保证其全面性和准确性,这些方面在今后的研究中还需进一步改进和完善。

参考文献:

[1] 韩宇,赵学文,李正风. 基础研究创新概念辨析及对相关问题的思考[J]. 中国基础科学,2001(3):35 – 40.

[2] 胡英奎,罗敏,王秀玲. 学术期刊编辑初审论文创新性的方法[J]. 编辑学报,2012(4):353 – 355.

[3] 盛杰. 期刊编辑对科技论文创新性的把握[J]. 编辑学报,2011(3):215 – 217.

[4] 朱大明. 科技期刊论文创新性鉴审的四个基本要素[J]. 科技管理研究,2011(9):199 – 201.

[5] 曹妍,朱瑞芳,韩世范. 应用德尔菲法构建护理论文创新性评价指标体系[J]. 护理研究,2017(17):2101 – 2103.

[6] AMJAD T, DING Y, DAUD A, et al. Topic-based heterogeneous rank[J]. Scientometrics, 2015, 104(1):1 – 22.

[7] HELLSTEN I, LAMBIOTTE R, SCHARNHORST A, et al. Self-citations, co-authorships and keywords: a new approach to scientists' field mobility? [J]. Scientometrics, 2007, 72(3): 469 – 486.

[8] LEE J Y. Exploring a researcher's personal research history through self-citation network and citation identity[J]. Journal of the Korean Society for Information Management, 2012, 29(1): 157 – 174.

[9] 魏瑞斌. 基于自引网络和内容分析的学者研究主题挖掘[J]. 情报学报,2015,34(6):635 – 645.

[10] HUMMON N, DEREIAN P. Connectivity in a citation network: the development of DNA theory [J]. Social networks, 1989, 11(1): 39 – 63.

[11] GOFFMAN W. Mathematical approach to the spread of scientific ideas--the history of mast cell research[J]. Nature, 1966, 212

(5061):449 - 452.
[12] JAHN M. Changes with growth of the scientific literature of two biomedical specialties[D]. Philadephia: Drexel University, 1972.
[13] 诺伊,巴塔盖尔吉. 蜘蛛:社会网络分析技术[M]. 林枫,译. 北京:世界图书出版公司,2012:328 - 331.
[14] 韩毅,金碧辉. 基于连通性的引文网络结构分析新视角:主路径分析[J]. 科学学研究,2012(11):1634 - 1640.
[15] 隗玲,方曙. 引文网络主路径研究进展评述及展望[J]. 情报理论与实践,2016(9):128 - 133.
[16] 宋歌. 学术创新的扩散过程研究[J]. 中国图书馆学报,2015(1):62 - 75.
[17] 韩毅,童迎,夏慧. 领域演化结构识别的主路径方法与高被引论文方法对比研究[J]. 图书情报工作,2013,57(3):11 - 16.
[18] 祝清松,冷伏海. 基于引文主路径文献共被引的主题演化分析[J]. 情报学报,2014,33(5):498 - 506.
[19] DENNIS A R, WIXOM B H, VANDENBERG R J. Understanding fit and appropriation effects in group support systems via meta-analysis[J]. MIS quarterly, 2001, 25(2):167 - 193.
[20] BROWN S A. Predicting collaboration technology use: integrating technology adoption and collaboration research[J]. Journal of management information systems, 2010, 27(2):9 - 54.
[21] DENNIS A R, FULLER R M, VALACICH J S. Media, tasks, and communication processes: a theory of media synchronicity[J]. MIS quarterly, 2008, 32(3):575 - 600.
[22] CRONIN B. Bibliometrics and beyond: some thoughts on Web-based citation analysis[J]. Journal of information science, 2001, 27(1):1 - 7.
[23] CRONIN B, SHAW D. Banking on different forms of symbolic capital[J]. Journal of the American Society for Information Science & Technology, 2010, 53(14):1267 - 1270.
[24] VAUGHAN L, SHAW D. Web citation data for impact assessment: a comparison of four science disciplines[J]. Journal of the American Society for Information Science & Technology, 2005, 56(10):1075 - 1087.
[25] 邱均平,余厚强. 替代计量学的提出过程与研究进展[J]. 图书情报工作,2013,57(19):5 - 12.
[26] LARIVIERE V, GINGRAS Y, SUGIMOTO C R, et al. Team size matters: collaboration and scientific impact since 1900[J]. Journal of the Association for Information Science & Technology,2015,66(7):1323 - 1332.
[27] PRATT J A, HAUSER K, SUGIMOTO C R. Defining the intellectual structure of information systems and related college of business disciplines: a bibliometric analysis[J]. Scientometrics,2012,93(2):279 - 304.
[28] SUGIMOTO C R, CRONIN B. Biobibliometric profiling: an examination of multifaceted approaches to scholarship[J]. Journal of the American Society for Information Science & Technology, 2014,63(3):450 - 468.
[29] SUGIMOTO C R, SUGIMOTO T J, TSOU A, et al. Age stratification and cohort effects in scholarly communication: a study of social sciences[J]. Scientometrics, 2016,109(2):997 - 1016.
[30] CRONIN B. Bibliometrics and beyond: some thoughts on Web-based citation analysis[J]. Journal of information science,2001,27(1):1 - 7.

Empirical Research on the Topic Innovation of the Paper Based on Self-citation Network and Main Path Analysis

Wei Ruibin
School of Management Science and Engineering, Anhui University of Finance and Economics, Bengbu 233030

Abstract: [**Purpose/significance**] This paper aims to study how to find the relative paper from large numbers of papers and study on the topic innovation of the paper. It may help researcher to read and use those papers. [**Method/process**] First, the paper defined the concept of the topic innovation and the self-citation network. And then the paper brought out the research idea, the tools and the method. At last the paper studied the Indiana University's innovation in the library and information science. [**Result/conclusion**] The result shows that the main path method can help to find these papers according to citation relationship from the self-citation network and analyze the topic innovation based on the result.

Keywords: topic innovation main path analysis self-citation network research topics